

Human capital and the quantity–quality trade-off during the demographic transition

Alan Fernihough¹

Published online: 22 September 2016

© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract This paper provides an empirical test of the child quantity–quality (QQ) trade-off predicted by unified growth theory. Using individual census returns from the 1911 Irish census, we examine whether children who attended school were from smaller families—as predicted by a standard QQ model. To measure causal effects, we use a selection of models robust to endogeneity concerns which we validate for this application using an Empirical Monte Carlo analysis. Our results show that a child remaining in school between the ages of 14 and 16 caused up to a 27 % reduction in fertility. Our results are robust to alternative estimation techniques with different modeling assumptions, sample selection, and alternative definitions of fertility. These findings highlight the importance of the demographic transition as a mechanism which underpinned the expansion in human capital witnessed in Western economies during the twentieth century.

Keywords Quantity–quality · Human capital · Demographic transition · Unified growth theory

JEL Classification J10 · N3 · O10

1 Introduction

Unified growth theory argues that the child quantity–quality (QQ) trade-off was a key mechanism in the transition from stagnation to modern economic growth (Galor 2011). Before the nineteenth century, economic growth was slow and fertility rates were persistently high. By the mid-twentieth century this situation had reversed in the Western world. The corre-

Electronic supplementary material The online version of this article (doi:[10.1007/s10887-016-9138-3](https://doi.org/10.1007/s10887-016-9138-3)) contains supplementary material, which is available to authorized users.

✉ Alan Fernihough
a.fernihough@qub.ac.uk

¹ Queen’s Management School, Queen’s University Belfast, Belfast BT9 5EE, UK

sponding increase in educational attainment observed during this period presents a narrative which is clearly consistent with various QQ theories. However, the findings in existing studies, which use post-demographic transition population samples, have been frustratingly inconclusive (see [Black et al. 2005](#); [Cáceres-Delpiano 2006](#); [Angrist et al. 2010](#)). Moreover, few microeconomic studies have tested the validity of this theory during this vital period.

The primary difference between this study and existing studies is that we use a large sample of individual level data from a historical population experiencing a demographic transition: 1911 Ireland. Ireland's relatively late fertility transition, alongside the availability of the completed individual census returns for the entire population, make it a viable testing ground for the importance of the QQ effect during the demographic transition.

Using the simple theoretical framework as in [Galor \(2011\)](#), we demonstrate how the child QQ model works, and assess the implications of this model for empirical research. This exercise demonstrates how the child QQ trade-off arises because parents substitute between fertility and child investments in response to relative price, income, and technological changes. It also implies that this substitution happens endogenously, with potential sources of endogeneity stemming from simultaneity, omitted variable bias, and measurement error.

Existing research has employed conventional instrumental variable (IV) estimators in an attempt to solve the endogeneity issue. However, IV estimators chiefly rely on valid exclusion restrictions, which given the highly endogenous nature of the child QQ model are extremely difficult to find in practice. Given the lack of a credible exclusion restriction, this study takes an alternative approach. Instead of using an exclusion restriction dependent IV approach, we adopt the methodology suggested in [Millimet and Tchernis \(2013\)](#), and estimate the QQ model using a number of estimators that either avoid or minimize endogeneity bias but do not require the researcher to make any unverifiable exclusion restriction assumptions. Much like cointegration, vector autoregression, or system/difference GMM, these estimators use alternative modeling assumptions but using cross-section data. These alternative modeling assumptions include substituting bias for variance in a manner similar to the conventional IV regression methodology and identifying information generated from heteroskedasticity in a first-stage modeling equation. Section 4 introduces these estimators in greater detail.

It is widely accepted that Ireland was something of an outlier in the European demographic transition ([Ó Gráda 1991](#)). Indeed, a casual look at early twentieth century Ireland suggests that the island's society was behaving in a way that contradicted the child QQ model. At this time the country had a relatively well-developed national education system, one that was founded in 1831, 39 years ahead of Britain. Illiteracy was rare amongst adults (see Table 1). However, the modernity suggested by Ireland's education system was untranslated to Irish demographic behavior. A simple comparison of [Ó Gráda \(1991\)](#) and [Wilson and Woods \(1991\)](#) suggests that marital fertility levels were over 1.6 times larger in 1911 Ireland relative to England in the same period. Nevertheless, [Ó Gráda \(1991\)](#) also showed how macro-level, aggregated statistics hid substantial heterogeneity in fertility behavior across Irish regions. When one takes a closer look at these data, it appears that fertility in the urban centers of Dublin and Belfast bore closer resemblance to the lower fertility patterns observed in more developed regions in Western Europe.

The demographic heterogeneity exhibited in Irish regions underlines an important point—aggregated data can obscure important individual-level relationships in any analysis. Our analysis shows how the resolution unit can be vital in QQ-type analyses. If we look

at aggregated county-level we find that counties with the lowest fertility also have the lowest school attendance. However, when we analyze more refined, higher resolution, data points, and focus on smaller geographic units inside these aforementioned counties, we find a negative conditional correlation connecting fertility to education—as suggested by the child QQ model. This finding underlines the potential for researchers to commit ecological fallacy when extrapolating results obtained via macro-level data to infer micro-level behavior and motivates the use of high-quality individual-level data in this paper.

This paper asks the following empirical question: do those attending school between the ages of 14 and 16 have fewer siblings, as predicted by the child QQ model? Schooling was effectively compulsory until the age of 14. If the QQ theory holds we expect that those who remain in school for longer to be from smaller families. OLS regressions on our large sample of individual-level data indicate that being in school reduces family level fertility by around 1 %. Interestingly, this effect intensifies as control variables accounting for potentially omitted variables, such as income and socioeconomic status, are included in the model specification. This suggests that any endogeneity bias may be biasing this coefficient towards zero.

To alleviate endogeneity concerns, we estimate a number of models proposed in [Millimet and Tchernis \(2013\)](#). The obvious criticism with these models is that whilst they avoid exclusion restrictions based assumptions, they require researchers to make alternative modeling assumptions. These alternative modeling assumptions either require the practitioner to correctly specify the regression equation or to use the identifying power generated by heteroskedasticity in a regression model of the endogenous treatment on the regressors. With these criticisms in mind, we probe the potential weakness of our empirical approach by conducting an Empirical Monte Carlo analysis as in [Huber et al. \(2013\)](#). This approach simulates “placebo treatments” using real data and thus assesses the performance of the exclusion restriction-free estimators given the causal effect is known. Assuming the presence of endogeneity, this analysis shows that the parametric version of the IV estimator proposed by [Klein and Vella \(2009\)](#) performs best in this application. This analysis also reveals a positive bias in the alternative exclusion-restriction free estimators proposed in [Millimet and Tchernis \(2013\)](#). Thus, we expect the results obtained from these estimators to underestimate (in absolute value) of the child QQ effect in our main empirical analysis.

Our Empirical Monte Carlo results help to contextualize our main results. Compared to the OLS regression results, the endogeneity bias-corrected models indicate a substantially larger QQ effect. The IV estimator based on [Klein and Vella \(2009\)](#), implies that a child staying in school between the ages of 14 and 16 reduces fertility substantially—by around 27 %. As predicted in the Empirical Monte Carlo results, the alternative estimators predict a smaller QQ effect. However, this effect, of around 10 %, is still significant. These results demonstrate the presence of a positive endogeneity bias. Our results are robust to different definitions of fertility (i.e. surviving of births), and potential issues of sample selection caused by maternal mortality.

Our findings illustrate the importance of human capital as a factor that underpinned demographic change in historical societies. These results appeal to the wider literature on long-run economic growth and technological progress ([Galor and Weil 2000](#); [Galor and Moav 2002](#)). Until recently, few research papers (notable exceptions include [Becker et al. \(2010\)](#) and [Klemp and Weisdorf \(2016\)](#)) have investigated this empirical link. The implication of our findings for the history of economic development is that the fertility transition was accelerated by parental investments in children, and therefore human capital, in Western economies in the late nineteenth and early twentieth centuries.

2 Theory and existing literature

Trends in economic growth and fertility during the period 1850–1950 presented a puzzle to economic demographers. Classical theory, typically attributed to Malthus, assumes that children are a normal good, and thus the income elasticity of parent's demand for children is positive. However, the emergence of modern economic growth alongside the demographic transition was at odds with classical theory. To answer this puzzle, [Becker and Lewis \(1973\)](#) proposed an extension to the classical economic model of fertility, and argued that parents made optimal (utility maximizing) child-rearing decisions across not one, but two dimensions: quantity and quality. Quantity refers to the number of children, whereas the meaning of quality can be loosely defined in terms of child-resources devoted human-capital augmenting education.

Since [Becker and Lewis \(1973\)](#)'s seminal contribution, the theoretical framework of the QQ model has been expanded, as in [Becker and Tomes \(1976\)](#). Several key contributions in the macroeconomic growth literature have cited this trade-off as the vital mechanism that fostered the emergence of sustained economic growth in Western economies. These citations include research by [Galor and Weil \(1999, 2000\)](#), and [Galor and Moav \(2002\)](#), who argued that an endogenous relationship between technological growth, the demand for human capital, and fertility emerged in the second phase of the Industrial Revolution. It is worth noting that a fundamental difference exists between the Becker-style QQ model and recent contributions ([Galor 2012](#)). In Becker's model, increased levels of income stimulate a decline in fertility via a substitution effect. Alternatively, parental preferences are such that the income elasticity with respect to quality is higher than with respect to quantity at higher levels of income. This is not the mechanism proposed in [Galor and Weil \(1999, 2000\)](#), and [Galor and Moav \(2002\)](#), who argued that the future return on their offspring's human capital caused parents to choose child quality over quantity. In essence, technological growth stimulated an economic expansion which drove an increase in the demand for human capital, that in-turn decreased fertility via the QQ trade-off.

As in [Galor and Moav \(2002\)](#), assume households maximize the following log-linear utility function:

$$u = (1 - \pi) \ln c + \pi [\ln n + \omega \ln h], \quad (1)$$

where n is the number of surviving children, h is the quality (a measure of human capital) of those children, and c represents all non-child consumption goods. The constant preference parameters $\pi \in (0, 1)$ and $\omega \in (0, 1)$ represent the household's preferences for children (over other forms of consumption) and child quality respectively. Each household is endowed with one unit of time, which they divide between labor market activities and child-rearing. There are two costs associated with child rearing, v^n is the fraction of a household's time dedicated to children ignoring child quality investments, while v^e is the cost (in time) associated with each unit of education e . If a household dedicates all its time to labor market activities it will generate an income y . Child-rearing represents time away from work, thus the price of each child is the opportunity cost associated with rearing. The budget constraint is therefore:

$$yn(v^n + v^e e) + c = y, \quad (2)$$

assuming the budget constraint binds. Additionally, it is assumed that human capital h is an increasing function of both education e and technological progress g :

$$h = f(e, g). \quad (3)$$

Taking Eqs. (1)–(3), the household's optimization implies the following Walrasian demand functions for fertility and child quality:

$$v^e f(\cdot) = \omega f_e(\cdot)(v^n + v^e e), \quad (4)$$

$$n = \frac{\pi}{v^n + v^e e}. \quad (5)$$

Equation (5) illustrates the negative relationship linking education and fertility, as $\partial n / \partial e < 0$. Additionally, we can see that the optimal level of education is determined both by child rearing and education prices alongside technological progress. The functional form of $f(e, g)$ is also a factor for determining the optimal level of education in Eq. (4).

The above model illustrates the basic mechanisms through which the child QQ model works and predicts why, at the micro level, families who choose to invest in their child's education have lower fertility. This model also highlights a number of econometric issues. Firstly, the family's education and fertility decisions are jointly determined. Multiple factors (prices, technology, and preferences) cause the QQ trade-off. This introduces the potential for simultaneity bias. Secondly, our empirical data does not include all of the relevant (and largely unobservable) variables in the QQ model. Instead, we use proxy information. For example, in our application we observed whether a child is in school and use this information as a proxy measure for education. This variable does not take into account the quality of school. Therefore, our data can be seen as a somewhat noisy representation of the true variable, and thus we suspect measurement error and attenuation bias. Thirdly, we expect there are exogenous factors simultaneously correlated with both child quantity and quality. This introduces the issue of omitted variable bias. In the Irish context urban status is negatively related to both fertility and education, a pattern that runs contrary to the QQ model. However, this effect stems from occupational differences specific to our data rather than the QQ model. In urban areas children are less likely to stay in school because greater opportunities for them to work outside the home exist.

The typical approach in empirical research has been to estimate one or both of the econometric models:

$$e_i = n_i \alpha_1 + \mathbf{X}_{1i} \boldsymbol{\beta}_1 + u_i \quad (6)$$

$$n_i = e_i \alpha_2 + \mathbf{X}_{2i} \boldsymbol{\beta}_2 + v_i \quad (7)$$

where the data variables contained in \mathbf{X}_{1i} and \mathbf{X}_{2i} consists of both control (the same variable is in both \mathbf{X}_{1i} and \mathbf{X}_{2i}) and instrumental (otherwise) variables. The variables e_i and n_i are as before and the coefficient parameters are: $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, α_1 , and α_2 . The u_i and v_i error terms measure unsystematic variation. Essentially, these are linear approximations to the structural QQ equations and the QQ effects are measured by the α_1 and α_2 parameters.¹

Estimating Eqs. (6) and/or (7) is complicated by endogeneity concerns, as previously discussed. Instrumental variables offer a popular solution to such concerns. In the context of the models expressed above this means including a variable in $\mathbf{X}_{1i}/\mathbf{X}_{2i}$ that is excluded from $\mathbf{X}_{2i}/\mathbf{X}_{1i}$. In this paper we are interested in estimating the causal effect of education on fertility. To estimate this QQ model using conventional IV via 2SLS or equivalent requires knowledge of a variable (or variables) that (monotonically) cause differences in education and only effects fertility through this education channel—the exclusion restriction.

Rosenzweig and Wolpin (1980) were the first to apply an exclusion restriction and IV methodology to this question. They argued that multiple births (like twins), for a given parity,

¹ The theoretical model in Eqs. (1)–(4) also suggest a nonlinear QQ effect. The nature of our data and empirical approach do not let me explore this empirically as in Mogstad and Wiswall (2009).

represent an exogenous increase in a family size. Using a sample of Indian households, their findings supported the QQ mechanism.

A multiple birth IV methodology was also performed in [Black et al. \(2005\)](#) who used a population sample of administrative records from Norway to estimate the effect of family size on school attainment. Interestingly the [Black et al. \(2005\)](#) findings suggest no relationship between a family's investment in education and fertility. A more recent study by [Angrist et al. \(2010\)](#) reached a similar conclusion based on Israeli data. In addition to multiple births, [Angrist et al. \(2010\)](#) used another IV based on the gender composition of the first and second born children. The validity of this instrument is justified on the basis of a parental preference for a mixed gender composition. [Cáceres-Delpiano \(2006\)](#) also used a multiple birth instrument to study the effects of family size using a sample of US children and found mixed support in favor of the QQ trade-off.

The validity of IV-based QQ estimates depends on the extent to which the exclusion restriction holds. [Rosenzweig and Zhang \(2009\)](#) question the use of multiple-birth instruments because twin births can cause an intra-household re-allocation of childhood resources. This intra-household re-allocation can obscure QQ effects, since parents appear to divert resources away from twin births to older siblings. [Galor \(2012\)](#) echoes these concerns, arguing that true tests of the QQ trade-off require either an exogenous change in the price of quantity or a change in the return to quality. Since multiple births impose a non-optimal level of household fertility, the response will not necessarily cause a family to deviate from the optimal level of quality. In essence, empirical testing of the QQ effect should examine how optimal levels of child quality respond to shocks in the exogenous variables.

Few studies have tested the QQ trade-off in historical populations. This is largely due to the scarcity of data containing the relevant empirical variables for a large number of observations. However, there are some exceptions. A study by [Bleakley and Lange \(2009\)](#) used the eradication of hookworm disease in Southern US states as a form of natural experiment. [Bleakley and Lange \(2009\)](#) argued that the eradication of this disease reduced the cost of child quality, and the subsequent increase in education and decrease in fertility were consistent with the QQ framework and also unified growth theory. [Hatton and Martin \(2010\)](#) use a unique individual level sample of British children in 1937–1939 to measure the relationship between family size and height, which serves as a proxy for human capital (health). Their results are also consistent with the QQ hypothesis as they show how family size was a key determinant of height. Another study by [Klemp and Weisdorf \(2016\)](#) looked at the relationship between fertility and literacy in historical England. Using exogenous variation caused by fecundity differences, [Klemp and Weisdorf \(2016\)](#) found that increases in sibship size caused reductions in adult literacy. Using aggregated regional data for mid-nineteenth century Prussia, [Becker et al. \(2010\)](#) approach the QQ trade-off from a macro perspective. Their analysis strongly supports the presence of a QQ trade-off. This macro-perspective finding also holds in nineteenth century France ([Murphy 2015](#)). The presence of an important child QQ effect is also supported in [Galor and Klemp \(2013\)](#). [Galor and Klemp \(2013\)](#)'s findings show that the process of natural selection appears to favor those with lower levels of fecundity in historical Quebec.

Given the the importance of the budget constraint in the child QQ trade-off the income-fertility relationship also holds relevance. There are a number of research papers that add context to the discussion of the economic determinants of historical fertility. [Clark and Hamilton \(2006\)](#) found a substantial fertility-income gradient in pre-industrial England. This finding was echoed in work by [Boberg-Fazlic et al. \(2011\)](#) who like [Clark and Hamilton \(2006\)](#) found a positive income-fertility relationship in pre-1800 England. [Cinnirella et al. \(2013\)](#) showed

how this mechanism operates in the short run, as families in pre-industrial England respond to short-term economic stress by increasing birth spacing thus lowering fertility.

In summary, there is a substantial literature on the child QQ trade off. However, this paper is unique as it estimated the child QQ-effect using a large sample of individual-level data in a historical context. Furthermore, this paper employs a series of econometric estimation procedures that are designed to alleviate the concern of endogeneity and thus does not rely on exclusion restriction based IV methods which have been questioned by others in this literature.

3 Data and context

Our study uses the individual level data returns from the 1911 Census of Ireland. The National Archives of Ireland provide full and unrestricted access to these returns. A key feature of these data is, as in Great Britain, all married women were asked how many children they had given birth to and also how many of these children had survived. The Registrar General was responsible for the collecting the census with the Irish police force served as enumerators. By 1911, the Irish police force had experience enumerating previous decennial censuses, alongside other enumeration duties like the collecting agricultural statistics (Guinnane et al. 2001). Completion of the census form was a simple procedure. Given the high level of literacy in Ireland at the time, it is reasonable to expect that most households contained at least one member capable of completing the enumeration form, although the use of these data entails some caveats. For example, whilst married women had little incentive to lie about child fertility or mortality, these data, like most self-reports, may, in some cases, contain some inaccuracies.

Lee (1969) illustrated how the Old Age Pensions Act of 1908 caused intentional age misreporting in 1911. Consequently, the age distribution for latter age-cohorts is skewed. However, this age-misreporting only occurred in older age cohorts (a large number of individuals claimed to be 73) and therefore is of little relevance to our research question. Guinnane et al. (2001) used these data to examine fertility behavior in Dublin city. As part of their analysis, Guinnane et al. (2001) examined age-misreporting by linking individual records in 1911 with early records in 1901. The results of this exercise indicated that limited age exaggeration occurred amongst women claiming to be under-50 in 1911. Furthermore, this analysis did not reveal any substantial misreporting biases after the socioeconomic stratification of individuals. Apart from the aforementioned age-misreporting, historical demographers regard these data as sufficiently accurate for research (Watterson 1988).

One previous study used the individual returns from the Irish census in their entirety (Fernihough et al. 2015), although research has been conducted on samples.² For example, Guinnane (1997) used a sample that linked the 1901 and 1911 censuses to look at family formation, the age of leaving home, and other demographic issues. Ó Gráda (2006) profiled demographic aspects of Dublin's Jewish community.

This paper uses school attendance as a measure of human capital. Educational attainment was an important determinant of social advancement in early twentieth century Ireland (Daly 1982). Entry into most clerical professions was contingent on school qualifications. Ireland's education system was comparatively quite advanced by 1911. Following the Irish Education Act of 1892, school fees for the majority of national/primary schools were abolished, while the

² A digitized version of the census was first released in August 2009 (see <http://www.census.nationalarchives.ie/about/index.html>).

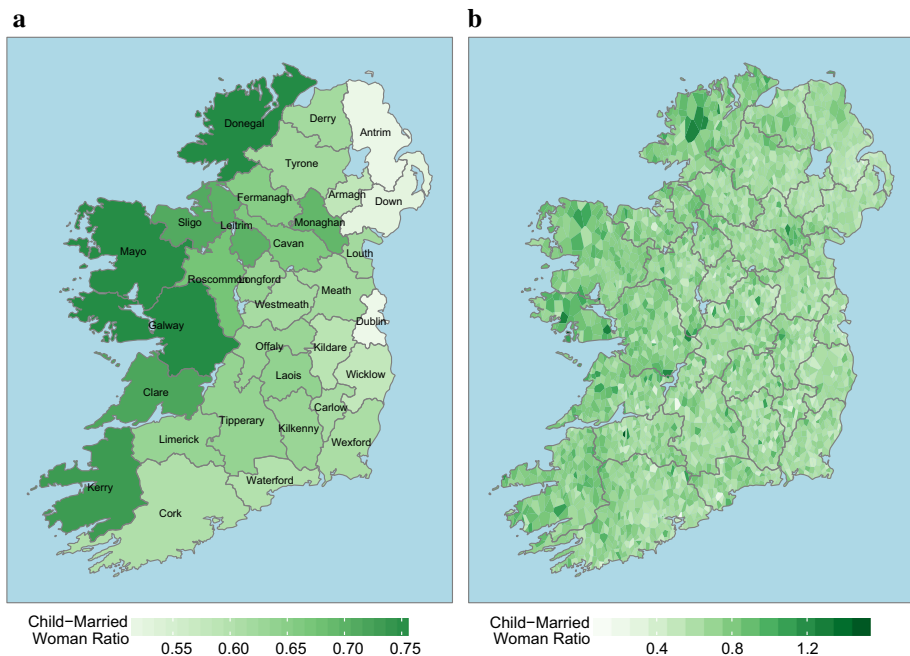


Fig. 1 Marital fertility, child-married woman ratio. **a** County-level resolution. **b** DED-level resolution

same act also introduced compulsory school attendance for all children between the ages of six and fourteen (Buachalla 1988). However, mandatory attendance could be circumvented for children aged twelve and older, provided they had found a source of regular paid employment (Patterson 1985).

The latter part of the nineteenth and early twentieth century saw a growth in secondary school attendance in Ireland. The introduction of state-organized exams in 1879 did much to foster this growth because a successful exam certificate was a valuable qualification for those wishing to join either the civil service or army (Coolahan 1981). A combination of both monetary and opportunity costs restricted the poorest from graduating to secondary level, although there were many exceptions. Religious bodies, particularly the Catholic Christian Brothers, built a substantial network of secondary schools enabling a large number of children from the poorer families to attain secondary-level education. There appeared to be significant returns to education in 1911 Ireland. However, this additional schooling should have implications for a family's budget constraint, as suggested by the child QQ model.

Figure 1 shows the spatial distribution of marital fertility in 1911 Ireland at two levels of resolution. Like Becker et al. (2010) we use the child-married woman ratio to measure marital fertility although these results are almost identical when we use child-woman ratios. This variable is defined as the number of children aged 0–4 divided by the number of married women aged 15–49. Plot (a) illustrates the spatial distribution of marital fertility aggregated at the county (32 national subdivisions) level. An east-west marital-fertility difference is apparent. This difference is underlined by the urban-rural split. The eastern seaboard contains large urban populations in the cities of Dublin and Belfast (nestled between counties Antrim and Down). The west of Ireland was very rural and hosted little economic activity outside the traditional agricultural sector in 1911. Guinnane (1997) showed that whilst rural fertility

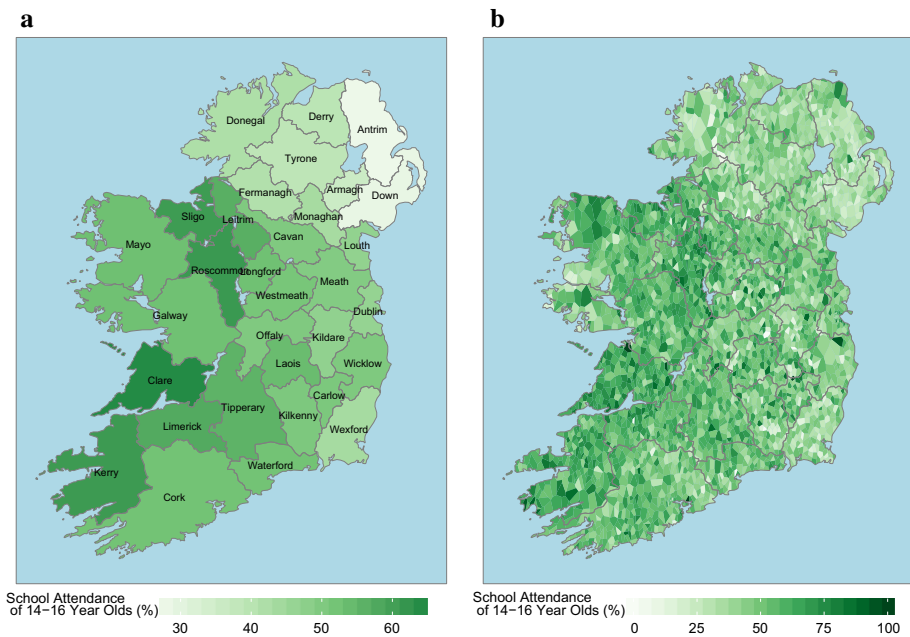


Fig. 2 School attendance for children aged 14–16 years. **a** County-level resolution. **b** DED-level resolution

did fall in response to the Great Irish Famine (1845–1852), this largely occurred through celibacy rather than within marriage. Nevertheless, Fig. 1 demonstrates that there must have been fertility control within marriage too, as we find large between county marital fertility differences, as in Ó Gráda (1991).

Plot (b) on the right-hand side of Fig. 1 shows marital fertility again, although aggregated at a much higher resolution. The regions here are the smaller district electoral divisions (DEDs). There are 3655 DEDs in these data. Overall, the marital fertility pattern similar to the county-level data shown in plot (a)—the areas with the highest levels of marital fertility are in the rural west and the lowest in the more urbanized east. However, a large degree of within-county heterogeneity is also evident. For example, large counties like Cork (in the south) have a mix of both high and low levels of marital fertility.

In this paper we use the individual returns to infer school attendance. The population were surveyed on their occupation and those attending school were typically enumerated using a description of “Scholar”. School attendance was effectively compulsory up to the age of 14, so we focus on school attendance amongst children between 14 and 16 years of age. Figure 2 repeats the analysis displayed in Fig. 1 with school attendance. Unlike, Fig. 1 the east-west dichotomy is not as evident. School attendance appears to be lower in the northern counties of plot (a). This is possibly reflects labor market conditions as, the more industrial northern counties would have had more opportunities for full-time employment outside seasonal agricultural employment. The opportunity costs associated with education were not constant across Irish counties.

The difference between plots (a) and (b) in Fig. 2 is also striking. The importance of within county heterogeneity was highlighted in Fig. 1 although it appears that this heterogeneity is even more prevalent in Fig. 2. Within counties some DEDs have high levels of school attendance whilst others do not. This potentially reflects the local supply of schooling.

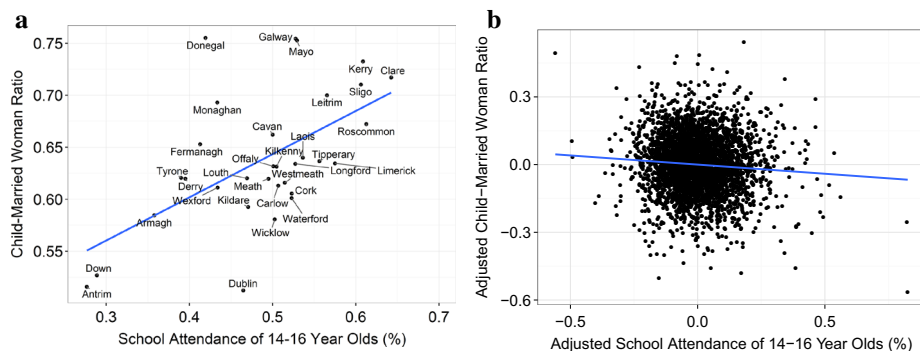


Fig. 3 Marital fertility and school attendance scatterplots. **a** County-level resolution. **b** DED-level resolution

Figure 3 illustrates the bivariate relationship between aggregated marital fertility and school attendance at both the county and DED level. Plot (a) shows the relationship at the county-level resolution. Contrary to the child QQ model, the correlation is positive in this plot. However, this relationship is reversed in plot (b) as we find a negative conditional relationship connecting fertility and education. This conditional relationship accounts for county-level fixed effects which may simultaneously affect fertility and education decisions.

The difference between plot (a) and (b) in Fig. 3 demonstrates the potential for researchers to commit ecological fallacy when trying to infer individual's behavior from aggregated data. For example, fertility in a county could be low and school attendance high, so the ecological inference here would be that families who send their children to school have less children. However, this assumes that a family from a county where fertility is low will also have low fertility—which might not be the case. The DED-level maps highlight a considerable degree of within-county heterogeneity. Therefore, it is perfectly plausible for the QQ relationship to hold at the individual level, but not the macro level. An example of this occurs in U.S. politics, where on average, wealthier people vote republican, but wealthier states vote democrat.

Previously [Brown and Guinnane \(2007\)](#) illustrated the importance of using disaggregated data in their study of the fertility transition in nineteenth century Germany and their findings appear to resonate here. Thus the ecological distinction motivates our use of individual-level data. These census data are cross-sectional, without retrospective information on either completed family size or education. However, since these data contain the entire population, the number of observations is large enough to make accurate inferences using this cross-section. Specifically, we evaluate whether school attendance is an important determinant of family size. School attendance contains students at both primary and secondary level as it was not uncommon for students to repeat their latter-primary years in order to obtain a more impressive school certificate ([Parkes 2010](#), p. 50). While we cannot observe the standard or level of schooling, it is reasonable approximation to assume that school attendance in these age-groups is an accurate indicator of family-level investment in education.

The observation unit in this analysis is all individuals between 14 and 16 years of age. The education variable is the aforementioned school attendance indicator. The census surveyed each married woman's fertility, asking for information on the number of births and surviving children. We have the entire census returns at our disposal and thus we match each 14–16 year old to the fertility information provided by their mother. The child QQ theory infers that parents investing more resources towards the human capital of their children causes decreases in family size. Lacking information on completed family size or education, this paper uses

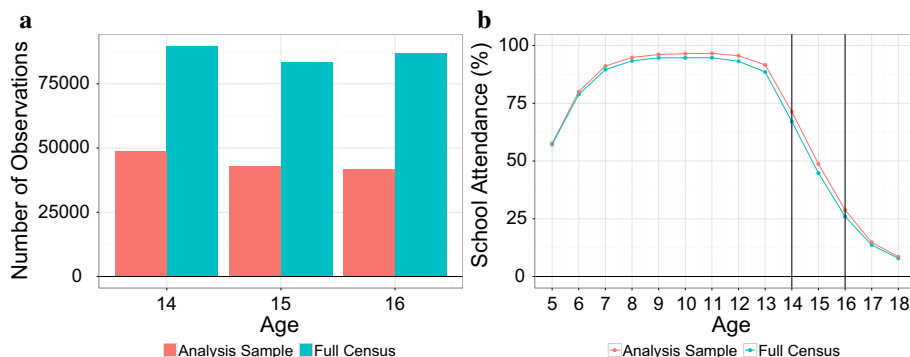


Fig. 4 Analysis sample comparisons. **a** Sample size comparisons. **b** School attendance and age

the census cross-section to take a snapshot of the Irish population in 1911. The dependent variable is fertility and the regressor of interest is a dummy variable indicating whether or not the observation is reported as a scholar.

In this model fertility is a function of schooling. However, the model is not temporally inconsistent because fertility is determined in the past and school attendance is recorded in the present. It is important to view the child's school attendance as an accumulation of human capital investments. For example, for a child to remain in school at age 14 they are highly likely to have been in school for several years previously. Thus their presence in schooling provides a suitable indicator for the household's investment in child quality over a long time frame. In the context of the child QQ model we see this as being a suitable variable for e_i —the time-invariant measure of education for an individual i .

The data we use contain a number of features which need to be processed prior to our econometric analysis. These census data represent one cross-section of the population at one specific point in time. Thus, we do not have information on completed schooling or fertility. A child in school at age 14 may have left a month after completing the census or remained in school for another 2 years. This is not information we have access to. The family-level fertility we observe relates only to one day in 1911. A mother reporting low fertility may go on to have more children than stated in the census. Also excluded are the children enumerated in residences where the mother is absent (through death or travel) and children with dead fathers, as the census only surveyed currently married women, not widows. We also omit a small number of families who live in multiple family households (as the family connections are two difficult to untangle) and children where the data are suspected to be inaccurate.³

The left-hand panel in Fig. 4 illustrates the difference in sample size between the full census and the observations used in our analysis. Most of the observations trimmed from the full census are individuals for whom we are unable to establish a maternal family link.⁴

³ For example, where the child is older than the number of years their parents have been married.

⁴ Initially, there were 260,340 children enumerated in the population. Requiring that these children be linked to both a father and mother present at the time of the census reduces the sample to 173,524 individuals or about 67 % of the initial population. Once we focus on all children who are born within 20 years of the marriage date the sample falls slightly to 151,172 (58 %). A number of observations had erroneous fertility data (i.e. the number of children alive exceeded the number born or the mother reported not giving birth to any children) and when we remove these the sample size is revised to 139,956 (54 %). Finally, our analysis sample does not include any observations with missing data on parental ages or occupations and where either or both of the parents are 30 years of age or younger. Omitting these observations from the analysis sample leaves us with 133,811 (51 %) data points.

Table 1 Summary statistics

Statistic	N	Mean	SD	Min	Max
Net fertility	133,811	6.446	2.487	1	18
Gross fertility	133,811	7.619	2.881	1	22
Scholar	133,811	0.507	0.500	0	1
Age	133,811	14.949	0.822	14	16
Mother's age	133,811	46.705	6.745	31	99
Father's age	133,811	52.324	8.866	31	100
Male	133,811	0.518	0.500	0	1
Literate mother	133,811	0.860	0.348	0	1
Literate father	133,811	0.827	0.378	0	1
Number of servants present	133,811	0.103	0.450	0	27
Belfast city	133,811	0.085	0.279	0	1
Dublin city	133,811	0.058	0.233	0	1
Religion: Roman Catholic	133,811	0.730	0.444	0	1
Religion: Church of Ireland	133,811	0.122	0.327	0	1
Religion: Presbyterian	133,811	0.111	0.314	0	1
Religion: Methodist	133,811	0.015	0.121	0	1
Religion: Jewish	133,811	0.002	0.040	0	1
Other religious affiliation	133,811	0.015	0.120	0	1
Uncategorized religious affiliation	133,811	0.006	0.075	0	1
Mother in labor force	133,811	0.062	0.240	0	1
Father in labor force	133,811	0.991	0.092	0	1
Father employed in agriculture	133,811	0.522	0.500	0	1
DED latitude	133,811	53.585	0.976	51.440	55.370
DED longitude	133,811	−7.480	1.240	−10.450	−5.462

The obvious concern with using a trimmed population sample is that this sample is unrepresentative of the population. However, as shown in the right panel, the % of children in school in our analysis sample is almost identical to the full census. Both follow the same decline trajectory between the ages of 14 and 16. We do not have any data on the family-level fertility of trimmed observations although these data are likely to be unrepresentative of fertility choice in the child QQ model anyway because they are from families for whom the completed family size choice has most likely been curtailed (due to the absence of a parent).

Table 1 contains the summary statistics for our data. We use two measures of fertility in this paper: net and gross. Net fertility is the logged number of surviving children reported by the mother in the family, whereas gross fertility is the logged total number of births. Our preferred outcome measure is net fertility, but we will also use gross fertility to demonstrate that our results are robust. We use logged measures so as to remove skewness from the distribution, but the results of the proceeding analysis are the same regardless of the unit of measurement. The average number of surviving children for each family is $\exp(1.771) \approx 5.877$. Roughly half of the children in the analysis sample are declared scholars in the census, which tallies well with the right-hand panel in Fig. 4. The remaining rows in Table 1 show the demographic and socioeconomic control variables used in this study. We are able to use information on

the gender of each child, their parent's literacy, whether they had servants in the house (an indicator for wealth), religious affiliation, and parental labor force information.

4 Empirical models

If the child QQ model holds, we expect within an economy each household's investment in education to be negatively correlated with their level of fertility. In our data, where individuals are aged between 14 and 16, those in school should be from households with lower levels of fertility. Consider the potential outcomes framework where $n_i(e)$ indicates the household fertility level n_i , which depends on a school "treatment" variable $e \in \{0, 1\}$. The causal effect of an individual being in school ($e = 1$) on fertility is thus:

$$\tau_i = n_i(e = 1) - n_i(e = 0), \quad (8)$$

the difference in potential outcomes. In this application we are interested in estimating the expected value of these potential outcomes or the average treatment effect (ATE):⁵

$$\tau_{ATE} = E[\tau_i] = E[n_i(e = 1) - n_i(e = 0)]. \quad (9)$$

In practice, we cannot perform the comparison in Eq. (9) as we only observe $e_i = 0$ or $e_i = 1$.

In our application we estimate the ATE ($\tau_{ATE} = \tau$) using the following system of equations:

$$n_i = \mathbf{X}_i \boldsymbol{\beta} + \tau e_i + v_i, \quad (10)$$

$$e_i^* = \mathbf{X}_i \boldsymbol{\gamma} + u_i, \quad (11)$$

$$e_i = \begin{cases} 0 & \text{if } e_i^* > 0 \\ 1 & \text{otherwise} \end{cases}, \quad (12)$$

where n_i and e_i are as before, \mathbf{X}_i is a vector of observable covariates, and v_i and u_i are two error terms. The conditional independence assumption (CIA) stipulates that the error terms are uncorrelated: $E(v_i u_i) = 0$. The CIA implies that if one was to estimate Eq. (10) via OLS the estimated parameter τ would not be subject to endogeneity bias. In the previous section we discussed potential sources of endogeneity in the QQ model and in practice we suspect that these error terms are related: $E(v_i u_i) \neq 0$. Consequently, we assume that these errors originate from the following bivariate normal distribution:

$$\begin{pmatrix} v_i \\ u_i \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_v^2 & \rho_{vu} \\ \cdot & 1 \end{pmatrix} \right), \quad (13)$$

where $\rho_{vu} \neq 0$, indicates endogeneity bias. The normalized inverse probability weighted (IPW) estimator of Hirano and Imbens (2001) offers an alternative to OLS regression. The advantage of propensity-score based estimators is that they are less reliant on the linear functional form. The IPW estimate of the ATE is given by the following formula:

$$\hat{\tau}_{IPW, ATE} = \left[\frac{\sum_{i=1}^N n_i s_i / \hat{P}(\mathbf{X}_i)}{\sum_{i=1}^N s_i / \hat{P}(\mathbf{X}_i)} \right] - \left[\frac{\sum_{i=1}^N (n_i (1 - s_i)) / (1 - \hat{P}(\mathbf{X}_i))}{\sum_{i=1}^N (1 - s_i) / (1 - \hat{P}(\mathbf{X}_i))} \right], \quad (14)$$

⁵ For reasons of parsimony we do not consider the average treatment on the treated (ATT) or average treatment on the untreated here.

where $\hat{P}(\mathbf{X}_i)$ represents the propensity score estimates obtained, in this application, from the predicted values of a probit regression fitted using the latent model expressed in Eqs. (11) and (12). Like OLS, the IPW estimator works under the assumption that the CIA holds.

Black and Smith (2004) recommend a trimmed sample version of the the IPW. Black and Smith (2004) derive a bias-minimizing propensity score for the ATT: $P(\mathbf{X}_i) = 0.5$. Based on this, the authors recommend a propensity based estimator with a thick support region where researchers only use data points for which: $P(\mathbf{X}_i) \in (0.33, 0.67)$. Millimet and Tchernis (2013) advance this methodology by deriving the bias minimizing propensity score for the ATE. Unlike the the ATT, the bias-minimizing propensity score of the ATE is is not fixed at $P(\mathbf{X}_i) = 0.5$. Instead it depends on the nature of the selection into the endogenous treatment (in our application schooling e_i). This bias is expressed:

$$B_{ATE}[P(\mathbf{X}_i)] = -(\lambda_0 + [1 - P(\mathbf{X}_i)](\lambda_1 - \lambda_0)) \left(\frac{\phi(\mathbf{X}_i\boldsymbol{\gamma})}{\Phi(\mathbf{X}_i\boldsymbol{\gamma})[1 - \Phi(\mathbf{X}_i\boldsymbol{\gamma})]} \right), \quad (15)$$

where ϕ and Φ represent both the probability and cumulative density functions of the standard normal distribution respectively, whilst the selection parameters λ_0 and λ_1 can be estimated from the following Heckman BVN selection model:

$$n_i = \mathbf{X}_i\boldsymbol{\beta} + \tau e_i + \lambda_0(1 - e_i) \left[\frac{\phi(\mathbf{X}_i\boldsymbol{\gamma})}{1 - \Phi(\mathbf{X}_i\boldsymbol{\gamma})} \right] + \lambda_1(e_i) \left[\frac{-\phi(\mathbf{X}_i\boldsymbol{\gamma})}{\Phi(\mathbf{X}_i\boldsymbol{\gamma})} \right] + v_i, \quad (16)$$

which in practice we estimate be replacing $\boldsymbol{\gamma}$ with sample estimates from a first-stage probit model. Note that the parameter τ in the regression model in Eq. (16) is the BVN estimate of the ATE parameter. If we replace the λ_0 , λ_1 , and $\boldsymbol{\gamma}$ parameters in Eq. (15) with their sample analogues, we can find the bias-minimizing propensity score (P^*), by performing a grid search.

Once P^* has been estimated this allows us to estimate a minimum biased (MB) version of the IPW:

$$\hat{\tau}_{MB,ATE} = \left[\frac{\sum_{i \in \Omega} n_i s_i / \hat{P}(\mathbf{X}_i)}{\sum_{i \in \Omega} s_i / \hat{P}(\mathbf{X}_i)} \right] - \left[\frac{\sum_{i \in \Omega} (n_i(1 - s_i)) / (1 - \hat{P}(\mathbf{X}_i))}{\sum_{i \in \Omega} (1 - s_i) / (1 - \hat{P}(\mathbf{X}_i))} \right], \quad (17)$$

where the sample Ω is a trimmed version of the full sample consisting of data points in some neighborhood around P^* : $(\underline{P}, \overline{P})$. Observations in this sample are removed based on their proximity to P^* , i.e. if $\hat{P}(\mathbf{X}_i)$ is either below \underline{P} or above \overline{P} . Observations further away from P^* are more likely to be discarded in the analysis. In essence, the MB estimator trades bias for variance. The amount of bias that needs to be traded off at the expense of variance requires the researcher to make a subjective decision about the “support region”. If the support region is too wide, the sample size will also remain large but this will limit the extent to which the bias is reduced. If the support region is too narrow, many observations will be discarded from the analysis and the estimated ATE will be less precisely measured and this will result in larger standard errors. Millimet and Tchernis (2013) recommend that \underline{P} and \overline{P} are chosen on the following basis:

$$\underline{P} = \max\{0.02, P^* - \kappa_\theta\} \quad (18)$$

$$\overline{P} = \min\{0.98, P^* + \kappa_\theta\} \quad (19)$$

where κ_θ is the smallest value such that at least $(\theta \times 100 \%)$ of the treatment and control groups are in the trimmed sample. The lower one is willing to set θ , the lower the bias MB estimate of the ATE, albeit at the expense of a higher sampling variance. Millimet and

Tchernis (2013) recommend setting $\theta \in \{0.05, 0.25\}$, and comparing, advice we adhere to in this paper.

Given that Eq. (15) allows us to estimate the endogeneity bias, we can use this information to create a series of bias corrected (BC) estimators. For example:

$$\hat{\tau}_{IPW-BC,ATE} = \hat{\tau}_{IPW,ATE} - \sum B_{ATE}[\hat{P}(\mathbf{X}_i)] \quad (20)$$

where $B_{ATE}[\hat{P}^*]$ is the sample estimate of the bias as in Eq. (15). This procedure can be also used to correct the MB estimators.

The final estimator we utilize is a parametric version of the IV estimator proposed in Klein and Vella (2009). This estimator exploits heteroskedasticity in the school treatment Eqs. (11) and (12) to estimate the causal impact of staying in school on fertility in Eq. (10). The rationale of this estimator is based on a wider literature that uses the heteroskedasticity for identification of endogenous regressor models (Klein and Vella 2010; Lewbel 2010; Hogan and Rigobon 2002). We reform the aforementioned probit model used to estimate γ , and instead estimate the following heteroskedastic probit model:

$$\Pr(e_i = 1 | \mathbf{X}_i) = \Phi \left(\frac{\mathbf{X}_i \gamma}{\exp(\mathbf{X}_i \delta)} \right) \quad (21)$$

as in Greene (2008, pp. 788–789). The maximum likelihood estimates in Eq. (22) $\hat{P}(\mathbf{X}_i)$ can then be used as a valid IV for e_i in Eq. (10), which we estimate using two-stage least squares.

Given the potential efficacy of the Klein and Vella (2009) estimator it is worth highlighting how it identifies the ATE parameter without an exclusion restriction. We have already seen, in Eq. (16), how the ATE can be estimated from Heckman's BVN model without an exclusion restriction. This is known as "identification by functional form" because it relies on the bivariate normal functional form being correct to achieve identification. This functional form-based identification also works in the Klein and Vella (2009) estimator when there is no heteroskedasticity in the education equation, i.e. $\exp(\mathbf{X}_i \delta) = 1$. The probit's nonlinear CDF means that the predicted values for education $\hat{P}(\mathbf{X}_i)$, will be linearly independent of \mathbf{X}_i and thus a valid IV for e_i . However, this assumption is very reliant on their being enough variation in the tails of the predicted $\hat{P}(\mathbf{X}_i)$ because the probabilities from $\Phi(\mathbf{X}_i \hat{\gamma})$ will be approximately linear around the center of this distribution.

The Klein and Vella (2009) estimator bypasses the aforementioned functional form assumption relying on variation in the tails. Instead it uses heteroskedasticity to achieve identification, which in econometric terms means that $Z_i \equiv [\mathbf{X}_i \gamma / \exp(\mathbf{X}_i \delta)]$ is linearly independent of \mathbf{X}_i . The nonlinearity generated by the normal CDF $\Phi(\cdot)$ is no longer a prerequisite for identification.

In practical terms, this methodology depends on three assumptions. Firstly, there must be sufficient heteroskedasticity in Eq. (22). If there is little or no heteroskedasticity $\hat{P}(\mathbf{X}_i)$ will be a weak IV for e_i , causing all the usual problems associated with weak IVs. Fortunately, this is a testable assumption. Secondly, we must assume that all of the variables in \mathbf{X}_i are exogenous. This is not a testable assumption, however one must also make this assumption when estimating an IV regression model with an exclusion restriction. Thirdly, we must also assume that the degree of endogeneity ρ_{vu} does not (conditionally) covary with any of the regressors in \mathbf{X}_i . Like the previous assumption, this is not one we can test but one that is already inherent in conventional IV strategies. In our application this assumption appears plausible because our preferred model specification includes a large set of control variables.

The econometric approaches outlined in the above are very appealing in situations where researchers do not have a valid exclusion restriction. Furthermore, these approaches could be

used in addition to the conventional IV approach as a robustness check when the instrument's validity is questionable (in a manner similar to over-identification procedures such as the Sargan test). These estimators may also appeal to applied empirical growth researchers even when the exclusion restriction is valid. Firstly, the instruments could be weak and thus the researchers may face a weak-IV problem. This problem has been acknowledged by [Bazzi and Clemens \(2013\)](#) who found that weak (and invalid) instruments are commonly used in the growth literature. Weak IVs subject the causal effect parameter to a greater standard errors and this uncertainty may make a researcher wary of exclusively using the conventional IV approach. Another well-known issue with the conventional IV approach is that it only identifies the local average treatment effect (LATE) ([Heckman and Urzúa 2010](#)). For historical growth studies with heterogeneous treatment effects this means that we might be estimating a local effect that differs substantially from the true parameter of interest ([Deaton 2010](#)). In other words, the IV methodology only provides us with an effect estimate that applies to a subsection of the population and if there are heterogeneous effects this means that our LATE differs from the ATE (which is typically the more important parameter of interest). Similarly, research that relies on exogenous variation created by historical events to instrument for contemporary endogenous factors (such as institutions) in growth regressions typically overestimate effect sizes because these instruments ignore persistence channels ([Casey and Klemp 2016](#)). The approach we follow allows us to avoid making any exclusion restriction assumptions that lack credibility.

5 Empirical Monte Carlo analysis

Table 2 lists all of the estimators described in Sect. 4. In textbook Monte Carlo analysis researchers simulate a variety of data generating processes (DGPs) and examine how effective certain econometric procedures are under different conditions. The drawback of this approach is that the entire DGP must be specified by the researchers. The Empirical Monte Carlo method, as developed in [Huber et al. \(2013\)](#), does not require researchers to simulate the full DGP for a model. Instead, it uses real data, in our case the census data, to simulate realistic “placebo treatments”, but leaving all of the other data variables unchanged. This approach is particularly useful in our application because we use estimators that rely on distributional assumptions in order to achieve identification.

The Empirical Monte Carlo approach works by simulating placebo treatments amongst the untreated. In our application we focus on the sample of children who were not in school in 1911. In Sect. 6 we discuss the model specification, but for now assume that we are using this paper's preferred specification of control variables: \mathbf{X}_i . As outlined in [Millimet and Tchernis \(2013\)](#), the Empirical Monte Carlo approach uses the following algorithm:

1. Using the full data sample estimate the probability of being in school $\hat{P}(\mathbf{X}_i)$ via the heteroskedastic probit model shown in Eq. (22).
2. With the sub-sample consisting of $e_i = 0$ designate placebo treatments as \tilde{e}_i :

$$\tilde{e}_i = I \left(\frac{\mathbf{X}_i \hat{\gamma}}{\exp(\mathbf{X}_i \hat{\delta})} + 0.3 + \zeta_i > 0 \right) \quad (22)$$

where $\zeta_i \sim \mathcal{N}(0, 1)$ is a random draw from a standard normal distribution. No relationship exists between the fertility outcome variable n_i and the placebo treatment \tilde{e}_i , so we expect the estimate of the ATE to be zero: $\tau_{ATE} = 0$. The error term and the outcome are uncorrelated $E[n_i \zeta_i] = 0$, so the CIA holds.

Table 2 Estimators employed

Name	Description
OLS	Conventional ordinary least squares regression
IPW	Normalized inverse probability weighted (IPW) estimator of Hirano and Imbens (2001)
MB-0.25	Minimum biased version of the IPW with $\theta = 0.25$, so at least 25 % of the treatment and control group must be in the trimmed analysis sample
MB-0.05	As above, but with $\theta = 0.05$
KV	Parametric Klein and Vella (2009) IV estimator
IPW-BC	IPW subtracting estimated bias
MB-0.25-BC	MB-0.25 subtracting estimated bias
MB-0.05-BC	MB-0.05 subtracting estimated bias
BVN	Heckman's bivariate normal model

Table 3 Empirical Monte Carlo results, root mean square errors

	OLS	IPW	MB-0.25	MB-0.05	KV	BC-IPW	MB-BC-0.25	MB-BC-0.05	BVN
CIA holds	0.004	0.004	0.007	0.013	0.021	0.088	0.085	0.087	0.088
CIA fails	0.029	0.030	0.027	0.028	0.025	0.092	0.089	0.091	0.092

These values are root mean squared errors based on 250 simulations of placebo “scholar” treatments for roughly 50 % of the control group. See the text for further details. Each model includes the full set of control variables, as detailed in the text, and county-level fixed effects. Lowest RMSE estimator given CIA assumption in bold

OLS ordinary least squares, IPW inverse propensity score weighting estimator, MB minimum-biased estimator using $\theta = 0.05$ or 0.25, KV Klein and Vella (2009) estimator, IPW-BC and MB-BC bias corrected versions of IPW and BC, BVN Heckman bivariate normal selection model

3. Use the estimators proposed in the Table 2 saving the estimates.
4. Repeat steps (1) to (3).

Given the algorithm above, the best estimator will be the one consistently closest to estimating an ATE of zero. We can also use the Empirical Monte Carlo approach to assess the performance of estimators when the CIA does not hold: $E[n_i \zeta_i] \neq 0$. To do this we draw the random error term in step 2 (ζ_i) from: $\zeta_i \sim \mathcal{N}(-0.04\tilde{n}_i, 1)$, where $\tilde{n}_i = (n_i - \mu_n)/\sigma_n$, and μ_n and σ_n are the mean and standard deviation of the fertility variable for the $e_i = 0$ sample. This setup creates an endogeneity bias of around -0.03 , a value similar to OLS estimate of the QQ effect as we will see in Sect. 6.

Table 3 details the root mean square errors (RMSEs) for our Empirical Monte Carlo analysis. We performed 250 repetitions of this placebo simulation procedure under the condition that the CIA both holds and fails. Figure 5 provides a visual representation of these results. When the CIA assumption holds, OLS has the lowest RMSE. This is followed by the IPW estimator. We can see this represented in Fig. 5, as the ATE estimates are centered around the null and show relatively little dispersion. The RMSE for the other estimators is larger. Looking at the top panel in Fig. 5 we can see how the MB and KV trade off bias for variance as these estimators are centered around $\tau_{ATE} = 0$, but with a greater amount of dispersion compared to OLS and IPW. The bias corrected and Heckman estimators (that rely exclusively on functional form assumptions) do not perform well when the CIA holds.

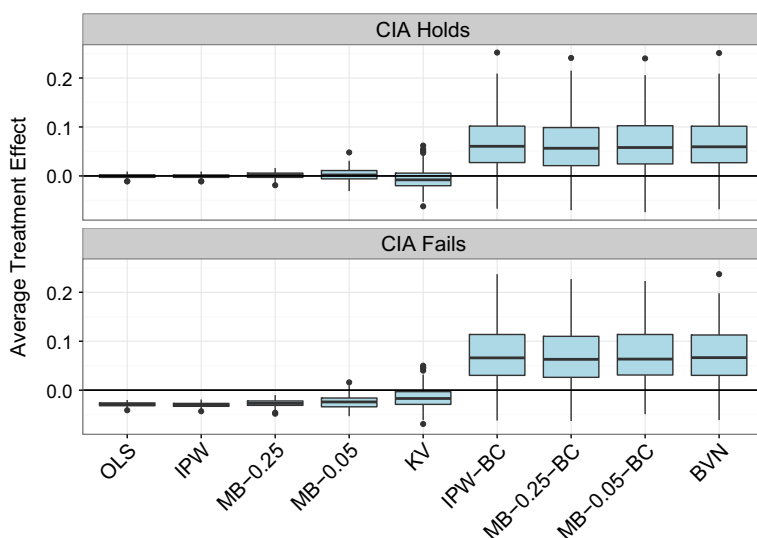


Fig. 5 Empirical Monte Carlo boxplots

This is either because the bivariate normality assumption is violated or there is insufficient nonlinearity in $\hat{P}(\mathbf{X}_i)$. Figure 5 depicts how these estimators are both biased and imprecise. Given the nature of these data, this bias is positive which goes against the child QQ model.

The second row in Table 3 repeats the Empirical Monte Carlo simulation analysis under the condition that the CIA fails. The magnitude of this bias is specified to replicate the OLS estimates we obtain in Sect. 6. In other words, if the true effect was zero but endogeneity created a negative bias of $\tau_{OLS,ATE} \approx -0.03$. The consequence of the CIA violation is evident in both Table 3 and Fig. 5. When the CIA fails, OLS is no longer the estimator with the lowest RMSE, the KV estimator is now the most accurate. As expected, the OLS and IPW estimators perform worse because they suffer from endogeneity bias. The bottom panel in Fig. 5 shows how both the OLS and IPW estimators are “precisely wrong” in the sense that they consistently estimate an ATE less than $\tau = 0$ with little variation. Table 3 demonstrates how the MB estimators are less biased, so have a lower RMSE, but the bias reduction comes at the expense of higher sampling variance, as seen in Fig. 5. This figure also illustrates how the minimum biased estimators are not unbiased, although a comparison between them and the OLS and IPW estimators is instructive as to the direction of this bias.

In this Empirical Monte Carlo setup the KV estimator is superior to the alternatives when the CIA fails. Table 3 shows how this estimator has the lowest RMSE whilst the bottom panel of Fig. 5 illustrates how, on average, the KV method exhibits considerably less bias than the alternate econometric procedures. As before, the BC and Heckman BVN methodologies have the least desirable properties. Much like when the CIA holds, when the CIA fails these methodologies still produce biased results. The bias associated with these simulations is positive. Given that we are using simulated placebo treatments on our actual data, this finding has implications for our empirical analysis. We expect the QQ effect to be positively biased when measured via Heckman BVN or the bias-corrected methods.

6 Results

The Empirical Monte Carlo results in Sect. 5 indicates the accuracy of the estimators outlined in Sect. 4 for the purpose of our empirical research question. Recall, we would like to estimate the parameter τ in the following linear model:

$$n_i = \mathbf{X}_i\boldsymbol{\beta} + \tau e_i + v_i \quad (23)$$

where τ is the ATE of a family letting a child remain in school on net fertility (number of surviving children as reported by the mother, n_i). A comparison of the estimated ATE for the full selection of estimation methodologies will be presented later. First, we use some standard OLS regressions with different control variables to estimate τ . These results are presented in Table 4. This table includes the coefficient estimates alongside their standard errors. The standard errors are (conservatively) clustered at the DED level to account for geographical/spatial autocorrelation.

The first column of Table 4 shows the relationship that connects the variable “Scholar” (whether the child is in school) to the net fertility in their family. In this model we include the following control variables: age dummy variables to capture the difference between 15 and 16 year old children compared to their 14 year old counterparts, a dummy variable to account for gender differences, and parental, both mother and father, age variables. This regression was estimated using 133,811 observations. The coefficient of the Scholar variable is -0.011 indicating that the choice for families to let their children stay in school beyond age 13 results in a 1.1 % drop in fertility.

Column (2) of Table 4 repeats the regression from column (1) adding in a county-level fixed effect. The counties are the 32 administrative districts in Ireland as shown in left-hand panels of Figs. 2 and 3. This is a potentially important control variable because it takes into account a number of geographical factors which may be simultaneously correlated with both school attendance and fertility, as we have previously demonstrated in Fig. 4. Including county fixed effects indicates a larger QQ effect than is shown in the first column. The results for this model imply that the school attendance causes a 3.6 % fall in fertility for this cohort.

A comparison of columns (1) and (2) highlights the importance of geographical controls. The results displayed in columns (3) and (4) show what happens when we include fixed effects that relate to smaller geographical units. In column (3) we include DED fixed effects, where the DED is a smaller geographic unit within counties. In column (4) we include street-level fixed effects, so the QQ effect found in this specification relies on within-street variation, a comparison of neighbors. This is appealing because street-level fixed effects account for all the unobserved heterogeneity, socioeconomic and cultural, that varies at the street level. For example, we would expect there to be less variation in wealth at a street level rather than at the wider county level. Interestingly, the inclusion of either DED or street-level fixed effects does not diminish the size of the QQ coefficient.

Column (5) contains the estimated coefficients and standard errors for a model specification that includes a more comprehensive set of control variables. These additional control variables account for nonlinearity in the parent’s age, religious affiliation (with Roman Catholic as the omitted category), dummy variables for Dublin and Belfast cities (in addition to county fixed effects), latitude and longitude co-ordinates, the number of domestic servants present (a sign of wealth), parental literacy indicators, and parental labor force attributes. Once again, the inclusion of additional geographic and socioeconomic controls has little impact on the QQ coefficient. The coefficient in column (5) implies that school attendance creates a 3.3 % drop in fertility. Finally, in column (6) we repeat the model specification from the previous column but include street-level fixed effects alongside fixed effects that account

Table 4 Net fertility regressions, OLS

	(1)	(2)	(3)	(4)	(5)	(6)
Scholar	−0.011*** (0.004)	−0.036*** (0.004)	−0.036*** (0.004)	−0.034*** (0.005)	−0.033*** (0.004)	−0.028*** (0.004)
Male	0.007*** (0.002)	0.005** (0.002)	0.004* (0.003)	0.004 (0.003)	0.004 (0.002)	0.003 (0.002)
Aged 15	0.002 (0.003)	−0.002 (0.003)	−0.000 (0.003)	0.000 (0.004)	−0.002 (0.003)	−0.001 (0.003)
Aged 16	0.007** (0.003)	0.001 (0.003)	0.003 (0.003)	−0.001 (0.004)	0.001 (0.003)	0.004 (0.003)
Mother's age	−0.007*** (0.000)	−0.008*** (0.000)	−0.008*** (0.000)	−0.008*** (0.000)	0.210*** (0.018)	0.212*** (0.018)
Father's age	0.009*** (0.000)	0.007*** (0.000)	0.006*** (0.000)	0.006*** (0.000)	0.024*** (0.002)	0.024*** (0.002)
Mother's age ² /100					−0.512*** (0.037)	−0.518*** (0.037)
Father's age ² /100					−0.069*** (0.003)	−0.069*** (0.003)
Mother's age ³ /1000					0.029*** (0.003)	0.029*** (0.002)
Mother's age × father's age					0.001*** (0.000)	0.001*** (0.000)
Literate mother					0.014*** (0.005)	0.016*** (0.005)
Literate father					0.023*** (0.004)	0.024*** (0.004)
Number of servants present					−0.076*** (0.005)	−0.066*** (0.005)
Religion: Church of Ireland					−0.079*** (0.006)	−0.070*** (0.006)
Religion: Presbyterian					−0.068*** (0.007)	−0.063*** (0.007)
Religion: Methodist					−0.088*** (0.017)	−0.074*** (0.016)
Religion: Jewish					0.159*** (0.028)	0.201*** (0.027)
Other religious affiliation					−0.049*** (0.012)	−0.044*** (0.012)
Uncategorized religious affiliation					−0.093*** (0.019)	−0.085*** (0.018)
Dublin city					−0.035** (0.015)	−0.034** (0.015)

Table 4 continued

	(1)	(2)	(3)	(4)	(5)	(6)
Belfast city					−0.083*** (0.012)	−0.078*** (0.011)
Mother in labor force					−0.031*** (0.007)	
Father in labor force					0.064*** (0.016)	
Father employed in agriculture					0.022*** (0.004)	
DED latitude					0.019* (0.010)	0.018* (0.010)
DED longitude					−0.052*** (0.006)	−0.052*** (0.006)
County FE	<i>N</i>	<i>Y</i>	<i>N</i>	<i>N</i>	<i>Y</i>	<i>Y</i>
DED FE	<i>N</i>	<i>N</i>	<i>Y</i>	<i>N</i>	<i>N</i>	<i>N</i>
Street FE	<i>N</i>	<i>N</i>	<i>N</i>	<i>Y</i>	<i>N</i>	<i>N</i>
Father's HISCO FE	<i>N</i>	<i>N</i>	<i>N</i>	<i>N</i>	<i>N</i>	<i>Y</i>
Mother's HISCO FE	<i>N</i>	<i>N</i>	<i>N</i>	<i>N</i>	<i>N</i>	<i>Y</i>
Observations	133,811	133,811	133,811	133,811	133,811	133,811
R ²	0.016	0.037	0.086	0.391	0.065	0.076
Adjusted R ²	0.016	0.037	0.060	0.171	0.064	0.072

The table presents coefficient estimates of an OLS regression of logged fertility (number of surviving children born to each observation's mother) on school attendance and various control variables. The intercept term is omitted from the above. Standard Errors are clustered at the DED level

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

for the full series of HISCO occupational codes associated with parental employment (van Leeuwen et al. 2002). However, the estimated QQ effect remains unchanged.

Table 4 consistently finds a QQ effect in the region of -0.035 . However, as discussed in Sect. 2, there are a number of sources of endogeneity which may render these OLS estimates invalid. In Sect. 4 we motivated the use of alternative models that estimate ATEs in the presence of endogeneity. Table 5 displays these results. Each column of Table 5 contains the average treatment effect estimate and empirical confidence intervals (ECI) obtained via cluster bootstrap for the 9 estimators discussed in Sect. 4 and labeled in Table 2. Note that we include all control variables listed in column (5) of Table 4 in the model design matrix \mathbf{X}_i . We use this comprehensive model specification as Millimet and Tchernis (2013) recommend that the model is over-rather than under-specified. However, our results are robust to the exclusion of multiple variables in \mathbf{X}_i .⁶

As we would expect, the OLS estimate of the ATE is the same as is shown in Table 4: -0.033 . In our Empirical Monte Carlo analysis the IPW method yielded estimates that were almost identical to the OLS values. This pattern is repeated in Table 4, although IPW is estimated with slightly greater precision as it has narrower ECIs. The Empirical Monte Carlo analysis showed that the MB estimators tend towards the true ATE as the choice of θ

⁶ Results from alternate model specifications are available upon request.

Table 5 Effect of school attendance on net fertility, average treatment effects

	OLS	IPW	MB-0.25	MB-0.05	KV	IPW-BC	MB-0.25-BC	MB-0.05-BC	BVN
ATE	-0.033	-0.032	-0.053	-0.077	-0.271	-0.097	-0.078	-0.102	-0.097
Lower 95 % ECI	-0.041	-0.040	-0.064	-0.098	-0.334	-0.184	-0.181	-0.192	-0.184
Upper 95 % ECI	-0.025	-0.023	-0.024	-0.020	-0.121	-0.017	-0.024	-0.020	-0.018

Treatment is defined as school attendance between the ages of 14 and 16. Net fertility is measured as the logged number of surviving children born to the mother of each child. Each model includes the full set of control variables, as detailed in the text, and county-level fixed effects. The 95 % empirical confidence intervals are obtained using 250 cluster-bootstrap repetitions where clustering is assumed to exist at the DED level

OLS ordinary least squares, IPW inverse propensity score weighting estimator, MB minimum-biased estimator using $\theta = 0.05$ or 0.25 , KV Klein and Vella (2009) estimator, IPW-BC and MB-BC bias corrected versions of IPW and BC, BVN Heckman bivariate normal selection model

is set to smaller values. Therefore, we can interpret the MB-0.25 and MB-0.05 estimates, -0.053 and -0.077 respectively, as indicating the presence of positive endogeneity bias in both OLS and IPW. Hence, the MB estimates suggest that the QQ effect is larger than we have previously estimated. Whilst MB estimators reduce bias at the expense of a higher ATE parameter variance, the 95 % empirical confidence intervals shown in columns (3) and (4) are sufficiently far from the null to show that their ATE achieves statistical significance at conventional p -value levels.

The KV estimator performed best when the CIA fails in our Empirical Monte Carlo analysis. Our previous results indicate that the CIA does not hold in this application, so we expect the KV estimator to be the most informative. The QQ effect measured by the KV model here is large: -0.271 , which implies that the choice for families to let their children stay in school beyond age 13 causes in a substantial 27.1 % decrease in fertility.⁷ This finding is consistent with the MB results which suggested that the OLS and IPW estimates are biased upwards. This bias could be due to measurement error or omitted variable bias since we would expect simultaneity to negatively bias QQ estimates.

The remaining columns of Table 5 display the results of the bias corrected (IPW-BC, MB-0.25-BC, and MB-0.05-BC) and Heckman BVN estimators. These estimators rely heavily on the joint normality assumption to achieve identification. Our Empirical Monte Carlo analysis showed how we would expect a violation of this assumption to yield positively biased ATEs. Considering that the Empirical Monte Carlo analysis uses our actual data, it has important implications for our application. The range of BC and BVN estimates, ranging from -0.078 to -0.102 , indicate a larger QQ effect. The Empirical Monte Carlo analysis showed that these estimators produced a bias of around 0.1. Applying this bias to the BC and BVN estimates here suggests a QQ effect in the region of -0.2 , a value similar in magnitude to that produced by the KV estimator.

The body of evidence presented in Table 5 points towards a large QQ effect. Our analysis proceeds by examining the robustness of this finding. One concern with the results shown in Table 5 is that they use “net fertility” as the outcome variable. This net fertility variable corresponds to the number of surviving infants in the household and one might be concerned with the potential for differential infant mortality to have an offsetting effect. Table 6 addresses this concern by using “gross fertility” as the outcome. This variable represents the number of children to whom the mother has given birth to, regardless of whether they survived or not. Additionally, we also include a variable that counts the number of children who died in \mathbf{X}_i , although the results are almost identical if this additional control is omitted. The results in Table 6 are very similar to those in Table 5.

These census data only capture a cross section of the county’s population at one period in time. Their ability to inform us of life-course demography is somewhat limited. For example, we do not know what the completed family size is for each family. Similarly, there may be an element in survivor bias in the sample of mothers aged over 40. Additionally, the effect of schooling on fertility might be conflated with birth order issues, a concern expressed in Black et al. (2005). To counter these potential issues, we repeat the analysis shown in Table 5 but for a trimmed sample consisting of those with mothers between 35 and 40 years of age and who are also the eldest in the family. This sample will largely consist of first-born children to the age cohort of mother’s less affected by survivor bias. The drawback of this approach is that we are now working with a much smaller sample size of 14,177. The results of this additional analysis are reported in Table 7 below.

⁷ To test the assumption of heteroskedasticity in the first-stage equation we performed a Wald test which comfortably rejected the null of no heteroskedasticity with $p < 0.001$. Full results available upon request.

Table 6 Effect of school attendance on gross fertility, average treatment effects

	OLS	IPW	MB-0.25	MB-0.05	KV	IPW-BC	MB-0.25-BC	MB-0.05-BC	BVN
ATE	-0.029	-0.028	-0.049	-0.073	-0.247	-0.093	-0.079	-0.103	-0.093
Lower 95 % ECI	-0.035	-0.035	-0.061	-0.095	-0.303	-0.167	-0.164	-0.182	-0.167
Upper 95 % ECI	-0.021	-0.020	-0.024	-0.018	-0.106	-0.022	-0.025	-0.018	-0.022

Treatment is defined as school attendance between the ages of 14 and 16. Gross fertility is measured as the logged number of children born to the mother of each child. Each model includes the full set of control variables, as detailed in the text, a variable measuring the number of children born to the mother who have died, and county-level fixed effects. The 95 % empirical confidence intervals are obtained using 250 cluster-bootstrap repetitions where clustering is assumed to exist at the DED level

OLS ordinary least squares, IPW inverse propensity score weighting estimator, MB minimum-biased estimator using $\theta = 0.05$ or 0.25 , KV Klein and Vella (2009) estimator, IPW-BC and MB-BC bias corrected versions of IPW and BC, BVN Heckman bivariate normal selection model

Table 7 Effect of school attendance on net fertility, average treatment effects, trimmed sample

	OLS	IPW	MB-0.25	MB-0.05	KV	IPW-BC	MB-0.25-BC	MB-0.05-BC	BVN
ATE	-0.071	-0.066	-0.067	-0.061	-0.216	-0.192	-0.177	-0.170	-0.194
Lower 95 % ECI	-0.088	-0.091	-0.114	-0.159	-0.364	-0.423	-0.411	-0.409	-0.423
Upper 95 % ECI	-0.056	-0.052	-0.042	0.005	-0.029	0.029	0.009	0.021	0.032

Treatment is defined as school attendance between the ages of 14 and 16. Net fertility is measured as the logged number of surviving children born to the mother of each child. Each model includes the full set of control variables, as detailed in the text, and county-level fixed effects. Trimmed sample refers to a sample consisting to only the first-borns to mothers aged between 35 and 40 ($N = 14,177$). The 95 % empirical confidence intervals are obtained using 250 cluster-bootstrap repetitions where clustering is assumed to exist at the DED level

OLS ordinary least squares, IPW inverse propensity score weighting estimator, MB minimum-biased estimator using $\theta = 0.05$ or 0.25 , KV Klein and Vella (2009) estimator, IPW-BC and MB-BC bias corrected versions of IPW and BC, BVN Heckman bivariate normal selection model

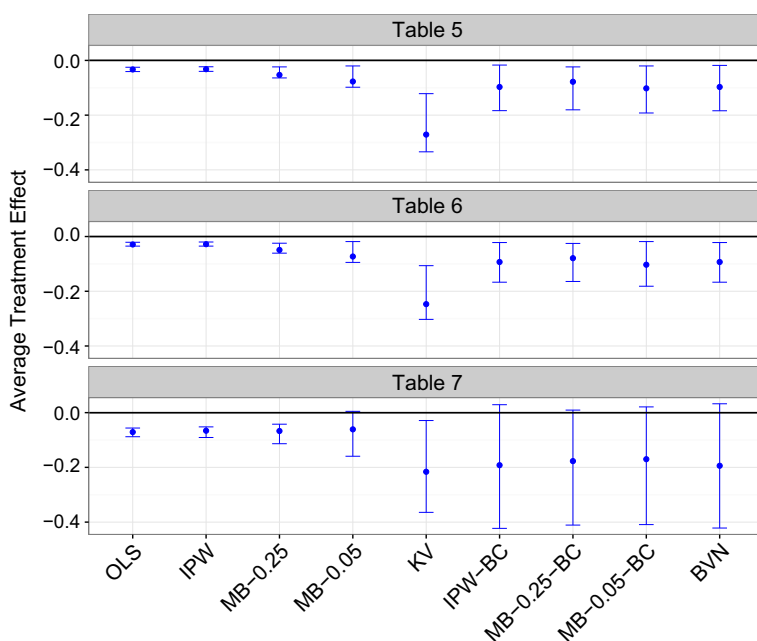


Fig. 6 ATEs from Tables 5, 6, and 7. The points in the above represent the ATE estimates and the error bars represents the 95 % empirical confidence intervals obtained using 250 cluster-bootstrap repetitions

The results in Table 7 once again indicate the existence of a substantial QQ effect. The estimated QQ effect here is, for the most part, slightly larger than was estimated in the previous tables. The OLS and IPW ATE estimates are over twice as large as their counterparts in the previous analysis tables. This indicates that there may have been some bias due to sample composition issues and/or a conflation with potential birth order effects. The magnitude of the KV ATE estimate does not change to reflect this, a result that signals this estimator's robustness. Figure 6 provides a graphical illustration of the model results shown in Tables 5, 6, and 7.

The body of evidence presented in the above unambiguously supports the presence of a QQ effect in 1911 Ireland. However, this evidence does not allow us to explore how this QQ effect might differ across the socioeconomic spectrum. In the following we examine the potential for heterogeneity in the QQ effect by re-running our analysis on three sub-samples stratified by socioeconomic status (SES) categories. We use the father's occupation to split these data. Fernihough et al. (2015) matched the 1911 Irish census occupation to HISCO codes and then to the HISCAM index of occupational social association (Lambert et al. 2013). The HISCAM index is a continuous variable that ranges from 28 (the lowest socioeconomic position) to 99. We first split these data by considering the Low SES group to be those who have a father who has an occupation below 58 on the HISCAM scale. This subsample consisted of 65,845 children with fathers typically employed as laborers, agricultural laborers, animal-drawn vehicle drivers, and other professions typically involving unskilled manual labor. Early twentieth century Ireland was primarily an agricultural society and this is reflected in the composition of the children's fathers occupations. A total of 57,035 of the observation's fathers were general farmers (although these data do not contain information on the size and quality of the family's farm). This broad category is designated with a HISCAM score of 58 and are categorized as our "Mid SES," group. In situations where the father's HISCAM score is above 58 we designate the observations as being "High SES." The High SES group

Table 8 Effect of school attendance on net fertility, average treatment effects by SES grouping

	OLS	IPW	MB-0.25	MB-0.05	KV	IPW-BC	MB-0.25-BC	MB-0.05-BC	BVN
Low SES									
ATE	-0.064	-0.066	-0.053	-0.048	-0.011	-0.040	-0.053	-0.048	-0.041
Lower 95 % ECI	-0.073	-0.076	-0.089	-0.109	-0.169	-0.134	-0.122	-0.133	-0.136
Upper 95 % ECI	-0.055	-0.056	-0.043	-0.020	0.055	0.050	0.042	0.049	0.051
Mid SES									
ATE	0.020	0.024	0.011	0.004	-0.144	-0.065	-0.064	-0.071	-0.067
Lower 95 % ECI	0.011	0.014	0.001	-0.020	-0.247	-0.223	-0.219	-0.223	-0.224
Upper 95 % ECI	0.032	0.0340	0.037	0.055	0.020	0.084	0.081	0.089	0.083
High SES									
ATE	-0.116	-0.090	-0.136	-0.102	-0.384	-0.378	-0.392	-0.358	-0.389
Lower 95 % ECI	-0.143	-0.125	-0.171	-0.250	-0.454	-0.591	-0.582	-0.593	-0.591
Upper 95 % ECI	-0.090	-0.049	-0.094	-0.021	-0.168	-0.098	-0.101	-0.087	-0.098

Treatment is defined as school attendance between the ages of 14 and 16. Net fertility is measured as the logged number of surviving children born to the mother of each child. Each model includes the full set of control variables, as detailed in the text (excluding the dummy variables that indicate whether the father's labor force and whether the father has a job in agriculture), and county-level fixed effects. Low SES is all observations where the father has a HISCAM index below 58, the mid SES sample consists of all observations where the father has a HISCAM index of 58 and the High SES group is all observations for whom the father had a HISCAM score above 58. All observations where the father did not have a HISCO score were dropped from this analysis. The 95 % empirical confidence intervals are obtained using 250 cluster-bootstrap repetitions where clustering is assumed to exist at the DED level

OLS ordinary least squares, IPW inverse propensity score weighting estimator, MB minimum-biased estimator using $\theta = 0.05$ or 0.25 , KV Klein and Vella (2009) estimator, IPW-BC and MB-BC bias corrected versions of IPW and BC, BVN Heckman bivariate normal selection model

is composed of 9,784 individuals living in households where the father worked in roles such as teaching or clerical administration or were a proprietor of their own business.

The results from our analysis run on the aforementioned sub-samples is displayed in Table 8. Overall, these results are somewhat mixed. There appears to be a substantial QQ effect amongst the Low SES grouping as the majority of these models estimate an effect ranging approximately from -4 to -6 %. An exception is the KV estimate, which estimates a QQ effect of -1.1 % albeit with wide ECIs. However, it should be noted that the identifying power generated by heteroskedasticity in the first stage is weaker here in comparison to other models.⁸ The QQ effect appears to be muted amongst the sons and daughters of general farmers, the Mid SES grouping. The results are inconclusive. The conventional OLS and IPW estimators suggest a small positive QQ effect whereas the MB estimators indicate that the OLS and IPW estimates are biased upwards, and the bias-corrected estimators all find a negative QQ effect. The results for the High SES group are unambiguous as a strong negative relationship is found in this subsample by all of the estimators.

Table 8 does not indicate the presence of a socioeconomic gradient in QQ effects. Instead, it provides evidence suggesting a weak QQ effect amongst families who operated as general farmers whilst a much stronger effect outside this large cohort. A speculative interpretation of Table 8 is that land-owning farming families faced different price parameters with respect to fertility and schooling because the children of these farmers could be employed in seasonal work that did not interfere with school attendance whilst families outside this social class perhaps found such opportunities, especially in urban areas, less feasible. The estimators in this paper trade off variance and bias, so once we start reducing the sample size, as in the Table 8 subsamples, the causal effect parameter estimates exhibit greater uncertainty. A consequence of this uncertainty is the wider confidence intervals, as displayed in Table 8 and these results must be treated with a degree of caution.

7 Conclusions

The emergence of unified growth theory has brought an imperative to understanding the historical relationship linking fertility and human capital. However, there is an absence of research examining whether the child QQ trade-off existed during the demographic transition.

To evaluate the child QQ trade-off in a historical society we used the complete census records from Ireland in 1911. If QQ theory holds, we expect those children who remain in school past the compulsory leaving age to be from families with lower fertility. This study highlights the importance of using disaggregated individual level data. There is potential for researchers to commit ecological fallacy when using marco/aggregated data to infer individual-level behavioral relationships. Our analysis finds a positive relationship connecting fertility and education at an aggregated level, however this relationship is reversed (as implied by the QQ model) once we focus on disaggregated units.

At the individual level we find a negative relationship between school attendance and literacy. The strength of this relationship increases as control variables are included in the model specification, suggesting the presence of a positive endogeneity bias. To formally account for this bias we use a series of estimators designed to estimate causal effects without exclusion restrictions. An Empirical Monte Carlo exercise that simulates placebo treatments reveals one estimator, a parametric version of the Klein and Vella (2009) IV procedure, performs very well in our application. The only additional assumption this IV methodology

⁸ The LR test for heteroskedasticity cannot reject the null of no heteroskedasticity in this model.

makes is the presence of heteroskedasticity in the first-stage equation, as opposed to an exclusion restriction—an assumption that is comfortably satisfied here.

Our preferred model specification, that of Klein and Vella (2009), estimates a child QQ effect of -0.266 . In our context, this means that the families chose to dedicate resources to their child's education by reducing their fertility by about 27 %. This substantial effect supports the use of the child QQ model to aid our understanding of simultaneous decline in fertility, increase in human capital, and subsequent economic growth of Western economies in the late-nineteenth and early-twentieth century.

Acknowledgements I acknowledge financial support from the Irish Research Council for the Humanities and Social Sciences. I am grateful to Cormac Ó Gráda, Morgan Kelly, Neil Cummins, Paul Devereux, Oded Galor, Jacob Weisdorf, Giampiero Marra, Chris L. Colvin, Alan De Bromhead, Mark E. McGovern, Daniel Millimet, four anonymous referees and participants at the Irish Quantitative History Group, the Antwerp FRESH meeting, and the 17th World Economic History Congress in Kyoto for their helpful comments for helpful comments and advice on an earlier or current draft of this paper. The usual disclaimer applies. Funding was provided by Irish Research Council for the Humanities and Social Sciences (IE).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Angrist, J. D., Lavy, V., & Schlosser, A. (2010). Multiple experiments for the causal link between the quantity and quality of children. *Journal of Labor Economics*, 28, 773–824.
- Bazzi, S., & Clemens, M. A. (2013). Blunt instruments: Avoiding common pitfalls in identifying the causes of economic growth. *American Economic Journal*, 5, 152–186.
- Becker, S. O., Cinnirella, F., & Woessmann, L. (2010). The trade-off between fertility and education: Evidence from before the demographic transition. *Journal of Economic Growth*, 15, 177–204.
- Becker, G. S., & Lewis, H. G. (1973). On the interaction between the quantity and quality of children. *Journal of Political Economy*, 81, 279–288.
- Becker, G. S., & Tomes, N. (1976). Child endowments and the quantity and quality of children. *Journal of Political Economy*, 84, 143–162.
- Black, S. E., Devereux, P. J., & Salvanes, K. G. (2005). The more the merrier? The effect of family size and birth order on children's education. *The Quarterly Journal of Economics*, 120, 669–700.
- Black, D. A., & Smith, J. A. (2004). How robust is the evidence on the effects of college quality? Evidence from matching. *Journal of Econometrics*, 121, 99–124.
- Bleakley, H., & Lange, F. (2009). Chronic disease burden and the interaction of education, fertility, and growth. *The Review of Economics and Statistics*, 91, 52–65.
- Boberg-Fazlic, N., Sharp, P., & Weisdorf, J. (2011). Survival of the richest? Social status, fertility and social mobility in England 1541–1824. *European Review of Economic History*, 15, 365–392.
- Brown, J. C., & Guinnane, T. W. (2007). Regions and time in the European fertility transition: Problems in the Princeton project's statistical methodology. *Economic History Review*, 60, 574–595.
- Buachalla, S. Ó. (1988). *Education policy in twentieth century Ireland*. Dublin: Wolfhound.
- Cáceres-Delpiano, J. (2006). The impacts of family size on investment in child quality. *Journal of Human Resources*, 41, 722–737.
- Casey, G., & Klemp, M. P. B. (2016). *Instrumental variables in the long run*. MPRA Paper 68696, University Library of Munich, Germany.
- Cinnirella, F., Klemp, M. P. B., & Weisdorf, J. L. (2013). *Malthus in the bedroom: Birth spacing as a preventive check mechanism in pre-modern England*. CAGE Online Working Paper Series, Competitive Advantage in the Global Economy (CAGE).
- Clark, G., & Hamilton, G. (2006). Survival of the richest: The Malthusian mechanism in pre-industrial England. *The Journal of Economic History*, 66, 707–736.
- Coolahan, J. (1981). *Irish education: Its history and structure*. Dublin: Institute of Public Administration.

- Daly, M. E. (1982). Social structure of the Dublin working class, 1871–1911. *Irish Historical Studies*, 23, 121–133.
- Deaton, A. (2010). Instruments, randomization, and learning about development. *Journal of Economic Literature*, 48, 424–455.
- Fernihough, A., Ó Gráda, C., & Walsh, B. M., (2015). Inter-marriage in a divided society: Ireland a century ago. *Explorations in Economic History*, 56, 1–14.
- Galor, O., & Klemp, M. (2013). *Be fruitful and multiply? Moderate fecundity and long-run reproductive success*. Working Papers 2013-10, Brown University, Department of Economics.
- Galor, O. (2011). *Unified growth theory*. Princeton, NJ: Princeton University Press.
- Galor, O. (2012). The demographic transition: Causes and consequences. *Cliometrica*, 6, 1–28.
- Galor, O., & Moav, O. (2002). Natural selection and the origin of economic growth. *The Quarterly Journal of Economics*, 117, 1133–1191.
- Galor, O., & Weil, D. N. (1999). From Malthusian stagnation to modern growth. *American Economic Review*, 89, 150–154.
- Galor, O., & Weil, D. N. (2000). Population, technology, and growth: From Malthusian stagnation to the demographic transition and beyond. *The American Economic Review*, 90, 806–828.
- Greene, W. H. (2008). *Econometric analysis* (6th ed.). Upper Saddle River, NJ: Pearson/Prentice Hall.
- Guinnane, T. W., Moehling, C., & Ó Gráda, C. (2001). *Fertility in South Dublin a century ago—A first look*. School of Economics Working Papers 200126, University College Dublin, Dublin Ireland.
- Guinnane, T. W. (1997). *The vanishing Irish: Households, migration, and the rural economy in Ireland, 1850–1914*. Princeton, NJ: Princeton University Press.
- Hatton, T. J., & Martin, R. M. (2010). The effects on stature of poverty, family size, and birth order: British children in the 1930s. *Oxford Economic Papers*, 62, 157–184.
- Heckman, J. J., & Urzúa, S. (2010). Comparing IV with structural models: What simple IV can and cannot identify. *Journal of Econometrics*, 156, 27–37.
- Hirano, K., & Imbens, G. (2001). Estimation of causal effects using propensity score weighting: An application to data on right heart catheterization. *Health Services and Outcome Research Methodology*, 2, 259–278.
- Hogan, V., & Rigobon R. (2002). *Using heteroscedasticity to estimate the returns to education*. NBER Working Papers 9145, National Bureau of Economic Research, Inc.
- Huber, M., Lechner, M., & Wunsch, C. (2013). The performance of estimators based on the propensity score. *Journal of Econometrics*, 175, 1–21.
- Klein, R., & Vella, F. (2009). A semiparametric model for binary response and continuous outcomes under index heteroscedasticity. *Journal of Applied Econometrics*, 24, 735–762.
- Klein, R., & Vella, F. (2010). Estimating a class of triangular simultaneous equations models without exclusion restrictions. *Journal of Econometrics*, 154, 154–164.
- Klemp, M., & Weisdorf, J. (2016). *Fecundity, fertility and the formation of human capital*. CAGE Online Working Paper Series 296, Competitive Advantage in the Global Economy (CAGE).
- Lambert, P. S., Zijedman, R. L., Van Leeuwen, M. H., Maas, I., & Prandy, K. (2013). The construction of HISCAM: A stratification scale based on social interactions for historical comparative research. *Historical Methods*, 46, 77–89.
- Lee, J. (1969). Irish agriculture. *Agricultural History Review*, 17, 64–76.
- Lewbel, A. (2010). Using heteroscedasticity to identify and estimate mis-measured and endogenous regressor models. *Journal of Business & Economic Statistics*, 30, 67–80.
- Millimet, D. L., & Tchernis, R. (2013). Estimation of treatment effects without an exclusion restriction: With an application to the analysis of the school breakfast program. *Journal of Applied Econometrics*, 28, 982–1017.
- Mogstad, M., & Wiswall, M. (2009). *How linear models can mask non-linear causal relationships. An application to family size and children's education (Discussion Papers 586)*. Research Department of Statistics Norway, Oslo, Norway.
- Murphy, T. (2015). Old habits die hard (sometimes). *Journal of Economic Growth*, 20, 177–222.
- Ó Gráda, C. (1991). New evidence on the fertility transition in Ireland. *Demography*, 28, 535–548.
- Ó Gráda, C. (2006). Dublin Jewish demography a century ago. *The Economic and Social Review*, 37, 123–147.
- Parkes, S. M. (2010). *A guide to sources for the history of Irish education, 1780–1922*. Dublin: Four Courts.
- Patterson, H. (1985). Industrial labour and the labour movement, 1820–1914. In L. Kennedy & P. Ollerenshaw (Eds.), *An economic history of Ulster, 1820–1940* (pp. 158–183). Manchester: Manchester University Press.
- Rosenzweig, M. R., & Wolpin, K. I. (1980). Testing the quantity–quality fertility model: The use of twins as a natural experiment. *Econometrica*, 48, 227–240.
- Rosenzweig, M. R., & Zhang, J. (2009). Do population control policies induce more human capital investment? Twins, birth weight and China's one-child policy. *Review of Economic Studies*, 76, 1149–1174.

- van Leeuwen, M., Maas, I., & Miles, A. (2002). *HISCO. Historical international standard classification of occupations*. Leuven: Leuven University Press.
- Watterson, P. A. (1988). Infant mortality by father's occupation from the 1911 census of England and Wales. *Demography*, 25, 289–306.
- Wilson, C., & Woods, R. (1991). Fertility in England: A long-term perspective. *Population Studies*, 45, 399–415.